Chapter 2
General Considerations in Clinical Research Design

Learning Objectives

1. The study population refers to all people who enter a study.
2. Common exclusion criteria in clinical studies include:
   a. Exclusion of people who have prevalent disease to focus on incident outcomes.
   b. Exclusion of people who have major disease risk factors to focus on the exposure of interest.
   c. Exclusion of people whose disease development may be missed in the study.
3. The choice of study population influences the generalizability of study findings.
4. The exposure is a factor that may explain or predict the presence of an outcome.
5. The outcome is a factor that is being explained or predicted in the study.
6. Observational studies observe the exposure; interventional studies assign the exposure.
7. Several factors favor causal inference in epidemiology research:
   a. Randomized evidence
   b. Strong associations
   c. Temporal relationship
   d. Exposure-varying response
   e. Biological plausibility

This chapter presents fundamental elements of a clinical/epidemiological research study: the study population, exposure, outcome, and the general study design. Specific study designs, along with their inherent strengths and weaknesses, are discussed in subsequent chapters. The chapter concludes with a discussion of factors that favor causal inference.
2.1 Study Population

2.1.1 Definition of the Study Population

The term study population, or patient population in a research study, refers to all of the people who enter a study, regardless of whether they are treated, exposed, develop the disease, or drop out after the study has begun. Typically, a study population originates from some larger source population, which is then narrowed using exclusion criteria.

Consider a study to address the hypothesis that estrogen use increases the risk of developing venous thromboembolism (VTE). Biological data suggest a link between estrogen use and VTE because estrogen interferes with a circulating factor that normally inhibits blood clot formation. One approach to studying this question in humans would be to identify a group of estrogen users and a group of nonusers, and then to follow them prospectively for the development of VTE. What exclusion criteria should be applied to best address the research question?

First, investigators may exclude women who have a previous history of VTE. Typically, studies of disease development focus on the incidence of disease, and therefore exclude people who have prevalent disease at the beginning of the study. Prevalent disease may be defined by any history of a chronic disease if that disease cannot be fully eradicated. For example, a previous history of coronary heart disease or diabetes is typically considered to represent “prevalent disease” in a clinical research study because these conditions are rarely cured, though frequently treated.

Second, investigators may exclude women with known major VTE risk factors, such as cancer or recent major surgery. Excluding women who have known causes of VTE will increase confidence that any new VTE cases that develop during the study can be attributed to the use versus nonuse of estrogen, rather than to some other factor. However, there are limits to using exclusion as a means to focus on a specific cause of disease. There are many risk factors for VTE, including genetic mutations, smoking, and kidney disease. Excluding women with any VTE risk factor would significantly diminish the size of the available study population, and would markedly restrict interpretation of study findings. In clinical practice, physicians do not test for a battery of rare mutations before prescribing estrogen. Results of a study that excluded women who had any predisposing mutation to VTE would have diminished generalizability to clinical practice.

Third, investigators may decide to exclude women whose VTE might be missed during the study. For example, subjects who plan on moving from the area may develop VTE in another geographic location and might not be counted. Subjects who have a history of frequently missing clinic appointments might be difficult to contact and less likely to complete surveillance procedures, potentially developing VTE that could be missed. When selecting a suitable study population, it is important for investigators to consider how they will capture the disease in question, and consider limiting the study population to subjects whose disease would be counted if it were to develop during the study.
2.1.2 Choice of Study Population and Generalizability of Study Findings

The choice of study population directly influences the generalizability or applicability of study findings. The following examples illustrate how different types of study populations influence the generalizability of the results.

Example 2.1. Clinic-based descriptive study of resistant soft tissue infections in children.

Study objective: Describe prevalence of resistant organisms in kids with soft tissue infection.

Study findings: Among 30 children with soft tissue infection demonstrated by wound culture, 7 (23%) had organisms that were resistant to first line antibiotics.

Study population: We studied 30 consecutive children with soft tissue infection from our outpatient pediatric clinic in greater Minneapolis. Children were included if they were 2–16 years old, had a soft tissue infection that required incision and drainage, and demonstrated organisms by culture examination.

Clinic-based studies such as these are generally the easiest and least expensive to conduct, because potential study subjects are often readily accessible to the investigators and important data may already be available as part of clinical practice. However, findings from these types of studies tend to be poorly generalizable to other populations. In this case, antibiotic resistance patterns may be specific to the geographic region where the study was conducted, and could be influenced by antibiotic prescription practices of this particular pediatric clinic. Results from this study will apply only to children who live in Minneapolis, and have the same age, socioeconomic background and pediatrician practice patterns as the children who attended this particular pediatric clinic. Moreover, clinic-based study populations tend to be relatively small and therefore highly subject to sampling variation or random fluctuation in results. Imagine that there were 500 total soft-tissue infections in this pediatric clinic and that 50 (10%) were caused by resistant organisms. Selecting random samples of 30 cases from the total group of 500 would yield a wide range of estimates of the proportion of resistant organisms.


Study objective: Estimate hip fracture rates in people with and without kidney disease.

Study findings: Late-stage kidney disease associated with fourfold greater risk of hip fracture.

Study population: The source population consisted of all male veterans with at least one outpatient primary care or internal medicine subspecialty clinic visit within the Northwest Veterans Integrated Service Network, a collection of eight Veterans...
Affairs facilities located in Washington State, Idaho, Oregon, and Alaska. Exclusion criteria were prior history of hip fracture, diagnosis of cancer, chronic dialysis, or renal transplantation.

Health network-based studies such as these offer an improvement in generalizability compared to clinic-based studies. In this case, the observed association of late-stage kidney disease with hip fracture is more broadly applicable to men in multiple geographic locations, and is not limited to the practice patterns of a single clinic or clinic. Further, the closed nature of the Veterans Affairs medical system increases the likelihood that hip fractures will be captured in this study population. This more general study population helps to support the hypothesis that late-stage kidney disease might play a causal role in the development of hip fracture. It is important however to note that this study population consisted predominantly of older men. Results may not apply to women, who have a considerably higher underlying prevalence of osteoporosis.

Example 2.3. Community-based study of lipoprotein a [Lp(a)] levels and stroke. Study objective: Examine association of Lp(a) with the risk of incident stroke in older adults.

Study findings: Higher Lp(a) associated with greater stroke risk in men, but not in women.

Study population: The Cardiovascular Health Study (CHS) is a community-based study of heart disease and stroke in 5,888 ambulatory adults aged 65 years and older. Participants were recruited from four communities by randomly sampling from age-stratified Medicare eligibility lists in each area. Subjects were excluded if they were institutionalized, were required a proxy to give consent, were required a wheelchair, or were receiving treatment for cancer.

Community-based studies such as these are typically the most costly and complex to conduct because they involve leaving the health care system in favor of the community for subject recruitment. The use of a community-based study population yields the most generalizable study findings, because many people in a community never see a doctor, let alone the inside of a hospital. Study findings for Lp(a) are expected to broadly apply to the general population of ambulatory older adults, not just those who are receiving health care.

2.1.3 Where to Find Information About the Study Population in a Clinical Research Article

Description of the study population is usually, but not always, described in the first one or two paragraphs of the methods section of a research article. This paragraph should explicitly define the source population from which study subjects were
selected and detail and justify the specific inclusion and exclusion criteria. This information may also be presented in flow chart form.

2.2 Exposure and Outcome

2.2.1 Definition

One broad segment of clinical/epidemiological research focuses on the relationship, or association, between an exposure and an outcome of interest. The term “exposure” is carried over from infectious disease epidemiology; however, it is used to describe any factor or characteristic that may explain or predict the presence of an outcome. Examples of exposures include the use of a particular medication, smoking, and blood pressure.

The term outcome refers to the factor that is being predicted. The outcome is often a disease, but can be any clinical characteristic, such as cholesterol level, vaccination status, or medication use.

The distinction between exposure and outcome is highlighted in the following examples.

Example 2.4. A study examines whether vaccination against pneumococcal pneumonia is effective at preventing the disease. Investigators review medical charts from 250 patients enrolled in a primary care clinic to determine whether they received the pneumococcal vaccine.

One possible study question might be:

\[
\text{Pneumococcal vaccine} \xrightarrow{\text{association?}} \text{Pneumococcal pneumonia}
\]

In this example, the exposure of interest is pneumococcal vaccination and the outcome of interest is pneumococcal pneumonia. The study would estimate the association of pneumococcal vaccination status with the risk of developing pneumococcal pneumonia.

Example 2.5. A study explores whether education level plays a role in a person’s decision to use herbal medications. A group of 500 people from a local shopping mall are asked to complete a questionnaire querying their education level and their frequency of herbal medication use.

A specific study question of interest might be:

\[
\text{Education level} \xrightarrow{\text{association?}} \text{Herbal medication use}
\]

In this example, the exposure of interest is education level and the outcome of interest is herbal medication use. Note that the outcome that is being predicted in this example is medication use. Other studies may examine medication use as the exposure, or predictor, of a disease.
Example 2.6. A study is conducted to examine whether heart failure influences survival after a first myocardial infarction. A total of 1,000 people who survive a first myocardial infarction undergo a history, physical examination, and echocardiography testing to determine the presence or absence of systolic heart failure. Subjects are followed until they die or drop out of the study.

A specific study question might be:

\[
\text{Systolic heart failure} \xrightarrow{\text{association?}} \text{Survival}
\]

In this example, the exposure of interest is systolic heart failure and the outcome of interest is survival. Note that the exposure in this case happens to be a disease. Other studies might focus on risk factors for systolic heart failure, and thus examine heart failure as the outcome of interest.

2.2.2 Specifying and Measuring the Exposure and Outcome

Effective clinical research requires highly focused definitions of exposure and outcome variables. For the pneumococcal vaccine study example, possible choices of specific exposures might be: (1) any previous pneumococcal vaccination (yes vs no), (2) recent pneumococcal vaccination within the last year (yes vs no), or (3) the number of years since last pneumococcal vaccination. Similarly, possible choices for the outcome variable, pneumococcal pneumonia, might be: (1) pneumonia, defined by chest x-ray findings or (2) pneumonia, defined by cough, fever, rales, and evidence of streptococcal DNA in sputum. For the heart failure example, the exposure variable, systolic heart failure, might be defined by a history of heart failure symptoms, such as shortness of breath and lower extremity swelling, plus evidence of a low cardiac ejection fraction measured by echocardiography.

Once exposures and outcomes are specifically defined, they should be measured as carefully as possible. For example, it is possible that medical chart records that document pneumococcal vaccination are less accurate than computerized pharmacy records that indicate actual disbursement of the vaccine. Errors in measuring exposure and outcome data are common; the consequences of measurement error are discussed in chapter 8.

2.2.3 Where to Find Exposure and Outcome Data in a Clinical Research Article

Information regarding ascertainment of exposure and outcome data is usually described beginning after the description of the study population in the methods section. This section should specifically define the exposure and outcome variables
and spell out exactly how they were collected and measured, so that the validity of
the study findings can be judged, and so that the study could be repeated under
similar conditions. If possible, studies should also describe the accuracy of the data
collection methods. For the pneumococcal vaccine example, the authors might state
in the methods section, “We defined pneumococcal pneumonia by a hospitalization
code for pneumonia plus culture evidence of streptococcus in the sputum. This defi-
nition correctly classified 88% of pneumococcal pneumonia cases compared to
gold-standard PCR testing for pneumococcus in a subsample of cases.”

2.3 Interventional Versus Observational Study Designs

Epidemiologic research studies can be broadly categorized as interventional or
observational. The distinction arises in the method by which study subjects are
exposed. An interventional study assigns exposure to study subjects, usually at random,
whereas an observational study observes the exposure, which occurs “naturally”.

For the previous example of estrogen use and venous thromboembolism, con-
sider first the interventional approach. At considerable effort and expense, investi-
gators could recruit a large group of postmenopausal women who were not
already using estrogen. Each recruited subject would then meet with a study
pharmacist who would flip a coin – if the coin comes up heads, the pharmacist
would assign estrogen therapy and if the coin comes up tails, the pharmacist
would assign an identical appearing pill that did not contain estrogen (placebo).
The coin flip would be conducted in secret, such that neither participant nor study
investigators were aware of the results. Following completion of this random
exposure assignment process, the baseline characteristics of exposed (estrogen
users) and unexposed (placebo) study subjects can be compared in a table of
baseline characteristics, which is usually the first table of a clinical research
article. Baseline characteristics of subjects assigned to estrogen versus placebo
are presented in Table 2.1.

Table 2.1 Baseline characteristics from a randomized trial comparing estrogen to placebo

<table>
<thead>
<tr>
<th></th>
<th>Estrogen use (N = 1814)</th>
<th>Placebo (N = 1814)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>61.9 (15.7)</td>
<td>61.9 (15.8)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>1241 (68.4)</td>
<td>1247 (68.7)</td>
</tr>
<tr>
<td>African–American</td>
<td>450 (24.8)</td>
<td>467 (25.7)</td>
</tr>
<tr>
<td>Other</td>
<td>123 (6.8)</td>
<td>100 (5.5)</td>
</tr>
<tr>
<td>Smoker</td>
<td>715 (39.4)</td>
<td>702 (38.7)</td>
</tr>
<tr>
<td>History of cardiovascular disease</td>
<td>795 (49.4)</td>
<td>833 (48.3)</td>
</tr>
<tr>
<td>Mean serum albumin (mg/dl)</td>
<td>3.92 (0.58)</td>
<td>3.90 (0.58)</td>
</tr>
<tr>
<td>Mean serum cholesterol (mg/dl)</td>
<td>188.6 (54.7)</td>
<td>188.4 (56.7)</td>
</tr>
</tbody>
</table>

All values in the table are either mean (standard deviation) or number of patients (percent).
Notice how the distributions of baseline characteristics are nearly identical between women assigned to estrogen versus those assigned to placebo, due to the random assignment of the exposure between groups. In contrast, one characteristic that is dramatically different between the groups is the use of estrogen at the beginning of the study; this should be 100% in the estrogen group and 0% in the placebo group. Assuming reasonable compliance with the assigned therapy during the study, potential differences in VTE outcomes between the two groups can be ascribed only to differences in estrogen use and not to other characteristics of estrogen users because in every other respect women assigned to estrogen are similar to those assigned to the placebo.

Now consider the observational approach. A population of postmenopausal women would be identified, but this time researchers would observe whether each study subject was using or not using estrogen. Investigators could ascertain estrogen status by querying computerized pharmacy records or by asking study participants to bring in their medication bottles to the study examination. Notice that in the observational study design, no participants receive a placebo. Baseline characteristics of estrogen users and nonusers are compared in Table 2.2.

Under the observational approach, exposure groups are more unbalanced with regard to some of their baseline characteristics. There are also a lower number of estrogen users; only 611 women in the observational study population were using estrogen. Some baseline characteristics appear to be unrelated to estrogen use, such as the serum albumin level. Others, such as previous cardiovascular disease, appear to be strongly linked with estrogen use, possibly because clinicians falsely believed that estrogen use reduced cardiovascular disease and may have prescribed estrogen for that purpose.

Like the interventional design, baseline estrogen use differs dramatically between exposure groups in the observational design: 100% versus 0%. However, if the two groups are found to have different VTE rates during follow-up, there will be residual uncertainty as to whether observed differences are due to estrogen use or due to differences in other characteristics of the women who used estrogen. This phenomenon is known as confounding and will be covered in detail in Chaps. 9 and 10. Freedom from confounding is the primary advantage of interventional trials over observational study designs.

<table>
<thead>
<tr>
<th>Table 2.2</th>
<th>Characteristics from observational study comparing estrogen use versus no use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estrogen users (N = 611)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>65.9 (15.7)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>528 (86.4)</td>
</tr>
<tr>
<td>African–American</td>
<td>66 (10.8)</td>
</tr>
<tr>
<td>Other</td>
<td>17 (2.8)</td>
</tr>
<tr>
<td>Smoker</td>
<td>177 (29.0)</td>
</tr>
<tr>
<td>History of cardiovascular disease</td>
<td>395 (64.6)</td>
</tr>
<tr>
<td>Mean serum albumin (mg/dl)</td>
<td>3.9 (0.58)</td>
</tr>
<tr>
<td>Mean serum cholesterol (mg/dl)</td>
<td>192.1 (54.7)</td>
</tr>
</tbody>
</table>
An important, but difficult concept is the distribution of unmeasured factors that are not presented in the baseline characteristics table, for example, exercise and dietary factors. In the interventional study design, if the sample size is reasonably large then random assignment will balance not only measured participant characteristics but also unmeasured characteristics that do not appear in the baseline table. So, it is expected that women assigned to estrogen in the randomized trial will have similar distributions of exercise patterns and dietary characteristics as those assigned to placebo. In contrast, there is no easy way to predict whether unmeasured characteristics will be balanced in an observational study and increasing the sample size will have no effect on this uncertainty.

2.4 Inferring Causation from Association Studies

2.4.1 Importance of Distinguishing Causation from Association

Epidemiologic research studies report associations between an exposure and an outcome, because association and not causation is actually observed. For example, studies detecting an association of LDL cholesterol levels with the occurrence of coronary heart disease did not observe LDL cholesterol entering arterial plaques and subsequently occluding blood vessels. This presumed sequence of events was based on multiple lines of evidence, from clinical association studies to basic experimental work. There is usually no way of directly observing an exposure causing a disease in humans.

Many associations are not causal. For example, receiving last rites in the intensive care unit is strongly associated with impending death; however, it is unlikely that receiving last rites causes a higher risk of death. Estrogen use has been associated with lower risks of heart disease in observational studies, but not in intervention trials, possibly because estrogen use is a marker of other healthy characteristics that are also linked with lower risks of heart disease, such as a healthier diet or compliance with other medical therapies. Separating association from causation in clinical research studies is critically important because uncovering causative factors in a disease process can lead to prevention and treatment. For example, an inference that the prone sleeping position caused sudden infant death syndrome (SIDS) led to successful prevention strategies that substantially lowered its risk. An inference that LDL cholesterol levels caused coronary heart disease led to the creation of specific drugs that lower LDL cholesterol, and to subsequent clinical trials proving their efficacy.

Inferring causality from epidemiological studies may not be easy. Causal inference is hampered in clinical research by the fact that (1) multiple exposures often influence a single disease outcome, for example, many people with low LDL cholesterol levels still develop heart disease, because they have other risk factors that play an important causal role and (2) many exposures take a long time to influence the outcome; for example, the effect of diet on the risk of cancer.
2.4.2 Factors Favoring an Inference of Causation

Although we can never be sure that a particular exposure causes a particular outcome, a number of factors can be used to help decide whether an exposure of interest is likely to be a cause of a disease, rather than just being associated with it.

2.4.2.1 Evidence Arising from Randomized Studies

Studies that randomly assign subjects to one treatment group versus another are generally the most powerful way to show that an exposure is a cause of an outcome. Large randomized trials are usually free from confounding, that is, characteristics of subjects assigned to one particular treatment are usually very similar, on average, to those of subjects assigned to another treatment. If outcomes differ between treatment groups, it is reasonable to conclude that the treatment is the cause of the difference. Unfortunately, randomized studies are limited to exposures that can be easily assigned to people, such as medications or devices.

2.4.2.2 Strength of Association

For both interventional and observational studies, a strong association between exposure and outcome increases the likelihood that the exposure is a cause of that outcome. Note that strength of association is not the same as statistical significance. For example, a case-control study observed that infants in the prone sleeping position had a fivefold greater risk of SIDS (odds ratio 5.0, \( p \)-value = 0.001). Although the \( p \)-value is important to rule out chance as a possible explanation for these findings, the strength of this association – that infants in the prone sleeping position were five times more likely to develop SIDS – was important for establishing the prone sleeping position as a cause of SIDS. One reason that strong associations often indicate causation is that there can only be so much bias and error in a well-conducted observational study. For the SIDS example, some errors in classifying SIDS versus other causes of death may have occurred, and there may have been other aspects of infants who sleep in the prone versus supine position that could have also influenced the risk of SIDS; however, it is unlikely that such potential errors would account for the entirety of such a strong observed association. In general, relative risks greater than 2.0 or less than 0.5 are considered to indicate strong associations.

2.4.2.3 Temporal Relationship Between Exposure and Outcome

For an exposure to be considered as a cause of a disease, there should be evidence that the exposure was present before the disease developed. Consider a study that examines the association between cyclophosphamide chemotherapy and the risk of
secondary bladder cancer. The study population includes patients who were free from bladder cancer at the beginning of the study, the exposure is the use of cyclophosphamide chemotherapy, and the outcome is newly diagnosed or incident bladder cancer that occurs at least 2 years after the initiation of chemotherapy. The investigators find that cyclophosphamide use is associated with a 4.5-fold greater risk of developing future bladder cancer. In this example, ensuring that the exposure (cyclophosphamide chemotherapy) clearly preceded the outcome (bladder cancer) in time strengthens the case for cyclophosphamide as a potential cause of secondary bladder cancer.

For second example, consider a hypothetical study that discovers high levels of a novel neurotransmitter, “DP-1” in the blood of people who have established major depressive disorder. These data alone do not clarify whether higher circulating DP-1 levels were present before the development of depression, or whether depression was present before DP-1 levels increased. The alternative possibility that DP-1 levels might rise in response to depression diminishes the case for causal inference.

2.4.2.4 Exposure-Varying Association

If a primary association between exposure and outcome is observed, the case for causal inference may be strengthened by additional evidence that the association differs predictably across different levels of the exposure. For the cyclophosphamide and bladder cancer example, the overall relative risk of secondary bladder cancer associated with cyclophosphamide use was 4.5. The investigators next examined the risk of bladder cancer associated with different cumulative doses of cyclophosphamide. Their findings are presented in Table 2.3.

<table>
<thead>
<tr>
<th>Cumulative cyclophosphamide dosage (g)</th>
<th>Relative risk of secondary cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Reference group</td>
</tr>
<tr>
<td>1–20</td>
<td>2.4</td>
</tr>
<tr>
<td>20–50</td>
<td>6.0</td>
</tr>
<tr>
<td>&gt;50</td>
<td>14.5</td>
</tr>
</tbody>
</table>

This stepwise increased risk of bladder cancer associated with each higher cyclophosphamide dosage strengthens evidence for a causal relationship between cyclophosphamide and bladder cancer. The “dose–response relationship” need not be limited to a medication, and can apply to different levels of any exposure. For example, childhood streptococcal infections have been associated with the development of neuropsychiatric syndromes, such as Tourette’s disorder. A well-conducted observational study observed that children who had a streptococcal infection were 2.2 times more likely to develop a future neuropsychiatric syndrome. The investigators strengthened the case for a causal relationship by further showing that the
risk of neuropsychiatric syndromes increased steadily with the number of previous streptococcal infections.

2.4.2.5 Biological Plausibility

Causal inference relies on translational and basic science knowledge to make sense of observed epidemiologic associations. Associations that have proven biological plausibility based on experimental data are more likely to be causal than those not supported by scientific evidence. Note that biological plausibility derives from scientific evidence obtained from other studies. For the example of LDL cholesterol levels and heart disease, multiple parallel studies took place: basic science studies demonstrated LDL cholesterol deposition in the arterial wall, translational studies showed enlargement of atherosclerotic plaque size by angiography among patients with higher LDL cholesterol levels, observational studies indicated associations of higher LDL cholesterol levels with a greater risk of developing clinical heart disease, and interventional trials demonstrated a reduced risk of death and cardiovascular disease in patients treated with drugs specifically designed to lower LDL cholesterol levels. This example highlights the importance of interdisciplinary collaboration for producing quality science and for moving the medical research field forward.
Epidemiology and Biostatistics
An Introduction to Clinical Research
Kestenbaum, B.
2009, XIV, 242 p. 46 illus., Softcover